*СТАТИСТИКА-

ЦЕ НАУКА , ЯКА ВИВЧАЄ, ОБРОБЛЯЄ Й АНАЛІЗУЄ КІЛЬКІСНІ ДАНІ ПРО НАЙРІЗНОМАНІТНІШІ МАСОВІ ЯВИЩА В ЖИТТІ.

* **МАТЕМАТИЧНА СТАТИСТИКА-**

ЦЕ РОЗДІЛ МАТЕМАТИКИ, ЯКИЙ ВИВЧАЄ МАТЕМАТИЧНІ МЕТОДИ ОБРОБКИ Й ВИКОРИСТАННЯ СТАТИСТИЧНИХ ДАНИХ ДЛЯ НАУКОВИХ І ПРАКТИЧНИХ ВИСНОВКІВ.

Місце математичної статистики в системі статистичних дисциплін

- Відомий статистик-математик професор Є.Є. Слуцький ще у 1912 році писав про значимість методів математичної статистики: "Ми приходимо таким чином до кардинальної вимоги, яку життя ставить діячам статистики: статистик повинен бути математиком, бо його наука є наука математична".

Місце математичної статистики в системі статистичних дисциплін

- В умовах широкого застосування методів сучасної математики в усіх галузях наукових досліджень, фундаментальних і прикладних, а також у вирішенні ряду практичних проблем суспільного життя увага надається математичній статистиці.

Місце математичної статистики в системі статистичних дисциплін

- У біології статистичні значення допомагають під час вивчення генетики, фізіології, екології. Нині жодна серйозна експериментальна робота з біології, медицини не обходиться без статистично обґрунтованого обсягу виконаних експериментів і довірчої оцінки отриманих результатів.

Висновки

- У даний час математична статистика знаходит широке застосування в економіці різних галузе народного господарства, біології, фізиці, хім медицині та ін. На основі її методів можна вирішуват і багато аналітичних задач в галузі економік Зокрема, кількісні характеристики, одержані результаті математико-статистичного аналіз дозволяють мати більш глибоке уявлення пр характер причинно-наслідкових зв'язків явищ, також одержати стійкі надійні параметри дл здійснення економічних розрахунків і особливо метою прогнозування.

* Задачі математичної статистики

1. Визначення закону розподілу випадкової величини за статистичними даними;

2. Перевірка правдоподібності гіпотез;

3. Знаходження невідомих параметрів розподілення.

* Основні терміни та поняття статистики:

Генеральна сукупність	Множина всіх можливих результатів спостереження (вимірювання)
Статистична вибірка, ряд	Множина результатів, які реально одержані в даному спостереженні (вимірюванні)
Варіанта	Одне із значень елементів вибірки
Варіаційний ряд	Упорядкована множина всіх варіант

Основні поняття математичної статистики з прикладами розв’язування задач

Приклад 1. За результатами контрольної роботи з теорії ймовірностей група за списком одержала такі оцінки:

8; 1; 3; 4; 9; 9; 10; 6; 8; 11; 7; 7; 8; 10; 6.

Варіаційний ряд має такий вигляд: 1, 3, 4, 6, 6, 7, 7, 8, 8, 8, 9, 9, 10, 10, 11.

* ЧИСЛОВІ ХАРАКТЕРИСТИКИ РЯДІВ ДАНИХ (вибірки)

РОЗМАХ ВИБІРКИ - різниця між найбільшим і найменшим значеннями величини у вибірці.

МОДА ВИБІРКИ - значення елемента вибірки, яке зустрічається частіше за інші.

МЕДІАНА ВИБІРКИ - це серединне значення впорядкованого ряду значень випадкової величини.

СЕРЕДНЄ ЗНАЧЕННЯ ВИБІРКИ - середнє арифметичне всіх чисел ряду даної вибірки.

* РОЗГЛЯНЕМО ЗАДАЧУ:

У відділі жіночого взуття протягом трьох днів було проведено обстеження для вивчення попиту на певні розміри взуття. За ці дні було продано 14 пар взуття 37, 38, 39, 40 і 41 розмірів.

1. Запишемо № розміру взуття і кількість відповідних йому проданих пар в таблицю:

* **Генеральна сукупність - кількість всіх пар взуття в магазині.**
* **Вибірка - 14 пар взуття.**
* **37, 37, 37, 38, 38, 38, 38, 38, 39, 39, 39, 40, 40, 41- статистичний, ранжований ряд.**

* **37, 37 , 37, 38, 38, 38, 38, 38, 39, 39, 39, 40, 40, 41 – вибірка**

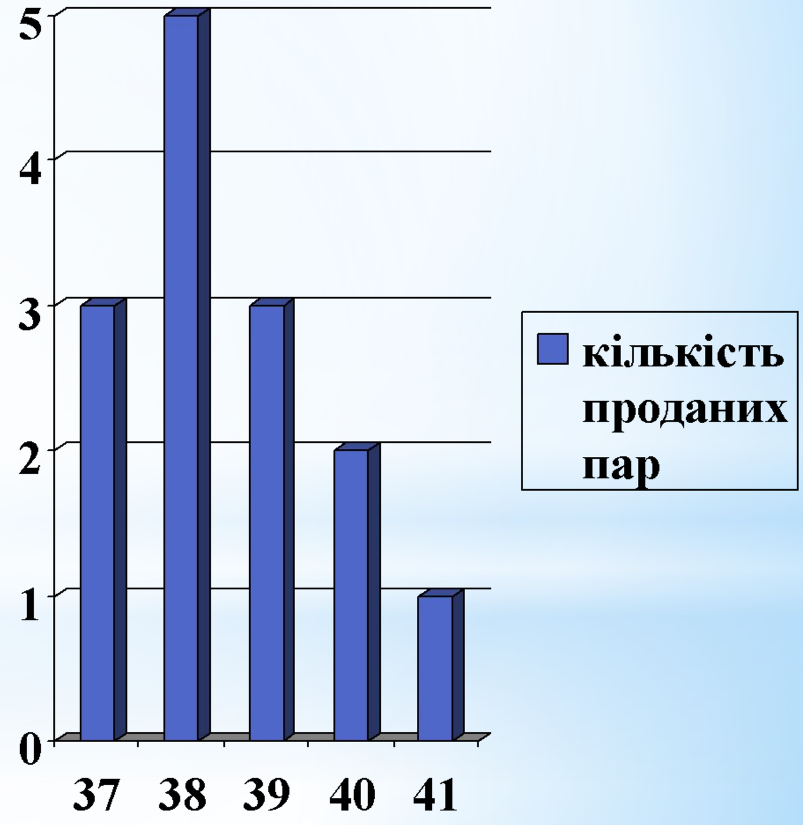
Розмір (варіанта)	37	38	39	40	41
Кількість проданих пар взуття (частота варіанти)	3	5	3	2	1
Відношення кількості проданих пар даного розміру до кількості всіх проданих пар (відносна частота варіанти)	3/14	5/14	3/14	2/14	1/14

- **Об'єм вибірки: 14.**
- **Розмах вибірки: 41-37=4.**
- **Мода вибірки: 38.**
- **Медіана вибірки: 38.**
- **Середнє значення :**

(37+37+37+38+38+38+38+38+
+39+39+39+40+40+41):14=38,5.

Розмір взуття	37	38	39	40	41	варіанти
Кількість проданих пар	3	5	3	2	1	частота варіант

- **Для наочного зображення даних спостереження можна використати стовпчикову діаграму. Також можна використовувати зображення у вигляді ламаної, яка називається полігоном частот.**
- **Графічні зображення дозволяють візуально охопити всю сукупність даних і скласти картину дослідження в цілому.**



Задача. Випадкова величина β- кут ковзання літака в момент скидання бомби. Проведено 20 бомбометань, в кожному зареєстровано кут ковзання β в тисячних долях радіана.

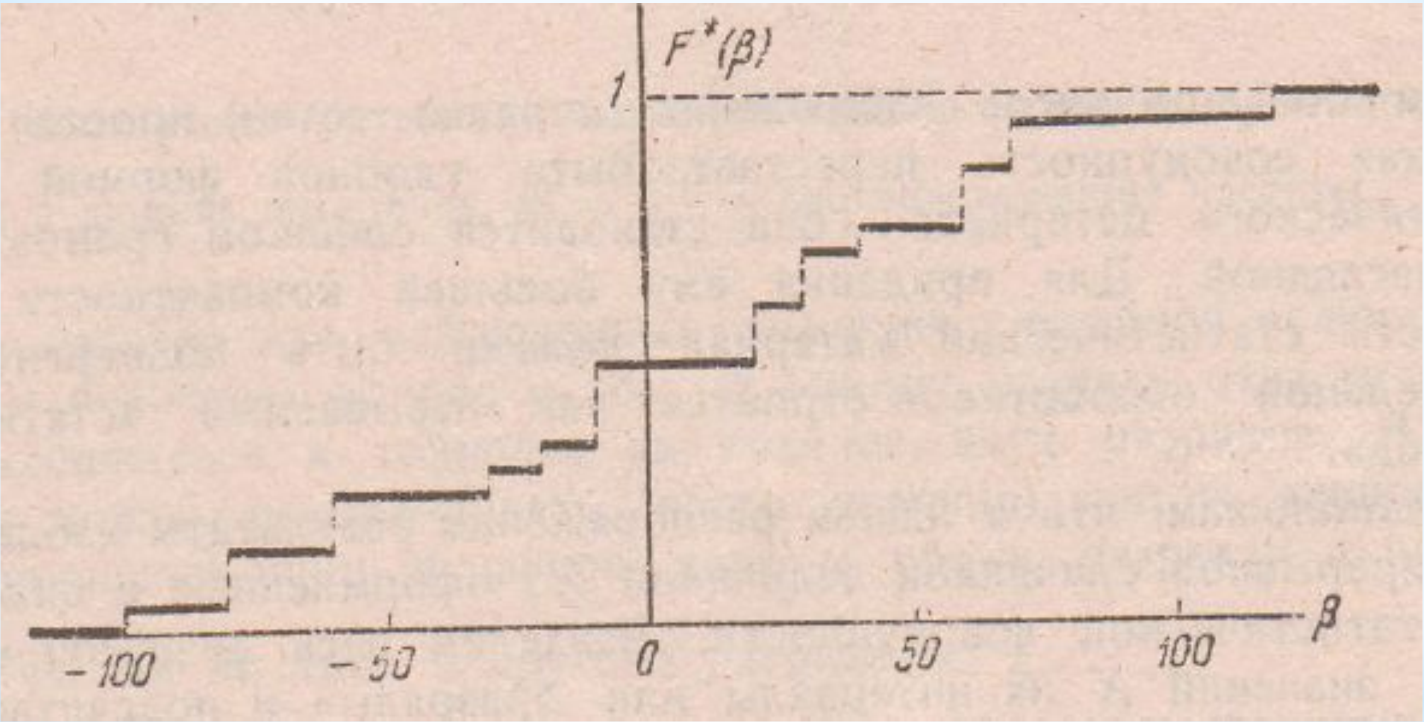
Результати зведені в простий статистичний ряд. Побудувати функцію розподілу.

i	β	i	β	i	β
1	-20	8	-30	15	-10
2	-60	9	120	16	20
3	-10	10	-100	17	30
4	30	11	-80	18	-80
5	60	12	20	19	60
6	70	13	40	20	70
7	-10	14	-60		

в тисячних долях радіана. Результати зведено в простий статистичний ряд. Побудувати функцію розподілу

$$F^*\left(x\right) = P^*\left(x \langle X\right)$$

Статистична функція розподілу



* Інтервальний статистичний ряд

Задача 15.85. Дані вимірювання частоти серцевих скорочень старшокласників подано у такому вигляді:

Частота скорочень x_i	65	69	70	71	72	73	74	75	76	77
Кількість учнів f_i	5	10	15	16	20	14	5	5	5	5

Потрібно:
а) побудувати графічне зображення даних таблиці;
б) згрупувати частоту серцевих скорочень у чотири групи і побудувати її інтервальний ряд розподілу; виконати відповідне графічне зображення.

Розв'язання: а) наведена таблиця є дискретним рядом розподілу частоти серцевих скорочень. Його графічним зображенням є полігон (рис. 15.12);

б) визначимо величину інтервалу: $x_{\max} = 77$, $x_{\min} = 65$, $l = \frac{77-65}{4} = \frac{12}{4} = 3$.

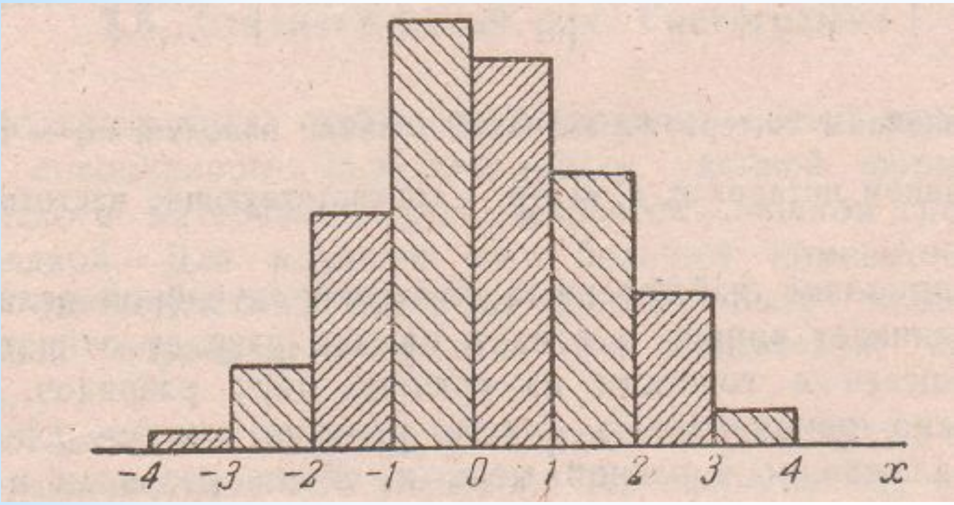
Будуємо інтервальний ряд розподілу:

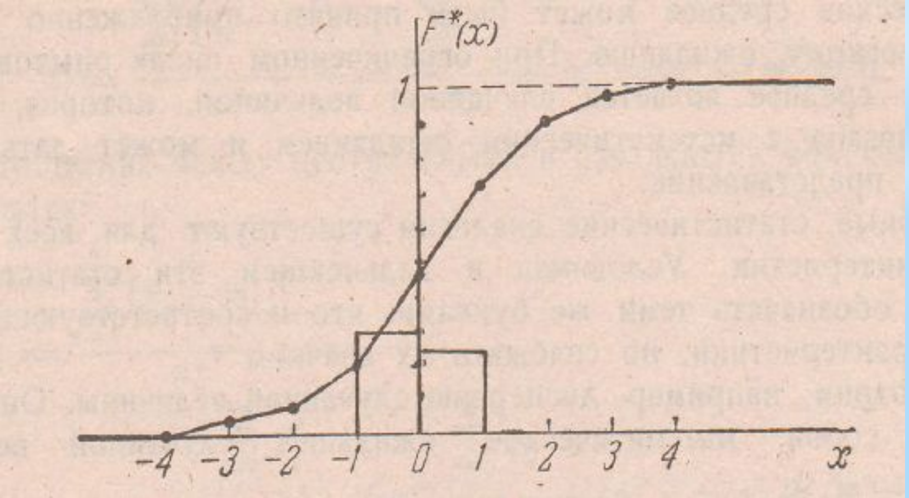
Частота скорочень I_i	65 – 68	68 – 71	71 – 74	74 – 77
Кількість учнів f_i	5	25	50	20

Задача 2. Проведено 500 вимірювань бокової похибки наводки при стрільбі з літака по наземній цілі. Результати вимірювань (в тисячних долях радіана) зведено в статистичний ряд. Побудувати гістограму, наближену статистичну функцію розподілу за даними статистичного ряду.

I_i	−4; −3	−3; −2	−2; −1	−1; 0	0, 1	1; 2	2; 3	3; 4
m_i	6	25	72	133	120	88	46	10
p_i^*	0,012	0,050	0,144	0,266	0,240	0,176	0,092	0,020

* Гістограма і статистична функція розподілення, побудовані за даними статистичного ряду



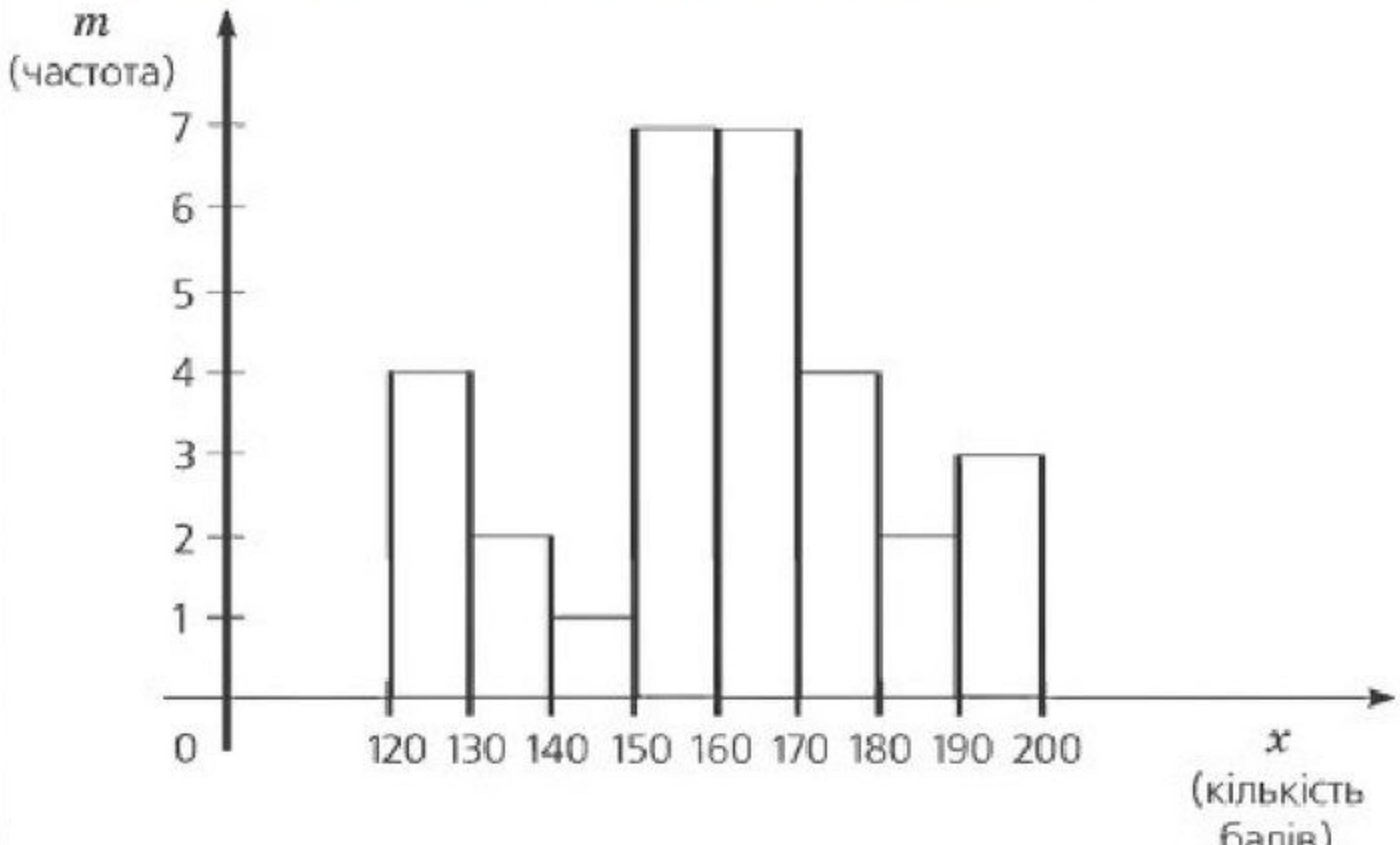


Приклад 3. Тридцятьма учасниками зовнішнього незалежно
оцінювання з математики були отримані такі бали:

120; 124; 125; 128; 135; 137; 142; 150; 151; 151; 151; 153; 155; 158; 16
160; 160; 165; 165; 166; 168; 170; 172; 175; 175; 180; 186; 190; 194; 200.

Гістограма частот:

Для прикладу 3 гістограма частот має такий вигляд:

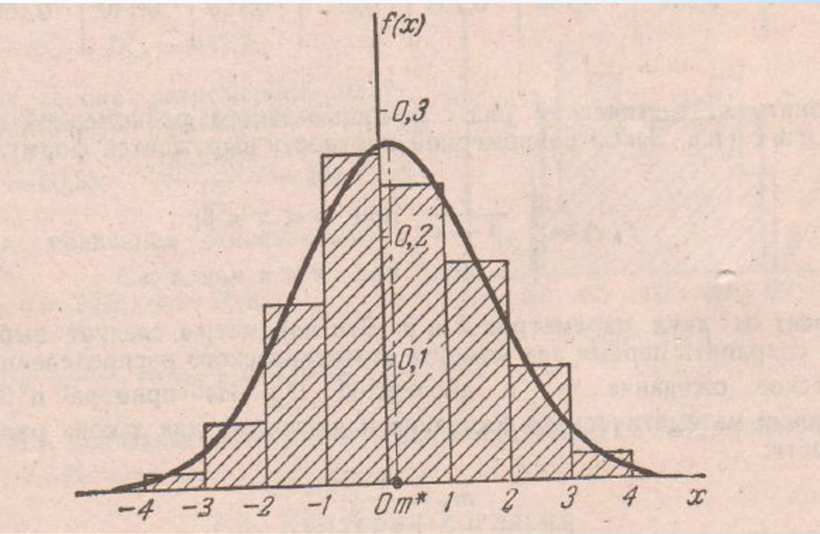


* Вирівнювання статистичного розподілу задачі 2 за допомогою нормального закону.

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}$$

$$m_x^* = -3{,}5 \cdot 0{,}012 - 2{,}5 \cdot 0{,}050 - 1{,}5 \cdot 0{,}144 - 0{,}5 \cdot 0{,}266 + 0{,}5 \cdot 0{,}240 + \\ + 1{,}5 \cdot 0{,}176 + 2{,}5 \cdot 0{,}092 + 3{,}5 \cdot 0{,}020 = 0{,}168.$$

$$\alpha_2^* = \sum_{i=1}^{\infty} \tilde{x}_i^2 p_i^* = 2{,}126.$$



$$D_x^* = \alpha_2^* - (m_x^*)^2 = 2{,}126 - 0{,}028 = 2{,}098.$$

$$m = 0{,}168; \quad \sigma = 1{,}448.$$

$$f(x) = \frac{1}{1{,}448\sqrt{2\pi}} e^{-\frac{(x-0{,}168)^2}{2\cdot 1{,}448^2}}$$

x	−4	−3	−2	−1	0	1	2	3	4
$f(x)$	0,004	0,025	0,090	0,199	0,274	0,234	0,124	0,041	0,008

* **Задача. З метою дослідження закону розподілу похибки вимірювання дальності за допомогою радіодальноміра зроблено 400 вимірювань. Результати досліду представлені у вигляді статистичного ряду. Вирівняти ряд за допомогою закону рівномірної густини**

I_i (м)	20; 30	30; 40	40; 50	50; 60	60; 70	70; 80	80; 90	90; 100
m_i	21	72	66	38	51	56	64	32
p_i^*	0,052	0,180	0,165	0,095	0,128	0,140	0,160	0,080

$$f(x)=\begin{cases} \dfrac{1}{\beta-\alpha} & \text{при} \quad \alpha<x<\beta; \\ 0 & \text{при} \quad x<\alpha \text{ или } x>\beta \end{cases}$$

$$m_x=\frac{\alpha+\beta}{2};$$

$$D_x=\frac{(\beta-\alpha)^2}{12}.$$

$\tilde{x}_i'$	—35	—25	—15	—5	5	15	25	35
p_i^*	0,052	0,180	0,165	0,095	0,128	0,140	0,160	0,080

$$m^*_{x'} = \sum_{i=1}^{k} \tilde{x}'_i p^*_i = 0{,}26.$$

$$\alpha^*_2 = \sum_{i=1}^{k} \left(\tilde{x}'_i\right)^2 p^*_i = 447{,}8,$$

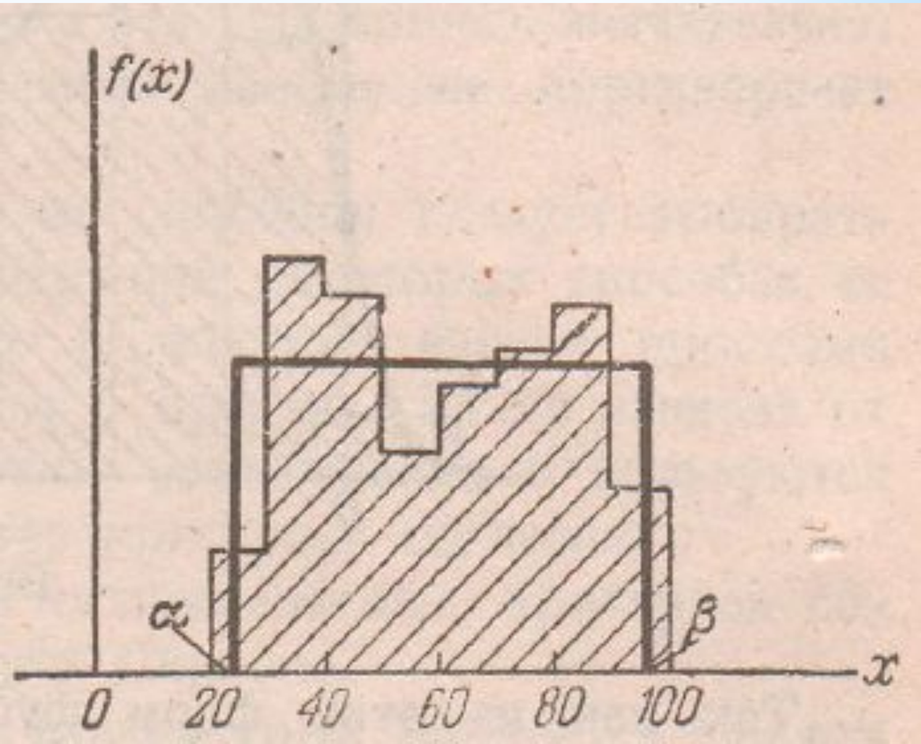
$$D^*_{x'} = \alpha^*_2 - \left(m^*_{x'}\right)^2 = 447{,}7.$$

$$m^*_x = m^*_{x'} + 60 = 60{,}26$$

$$\frac{\alpha+\beta}{2} = 60{,}26; \qquad \frac{(\beta-\alpha)^2}{12} = 447{,}7.$$

$$\alpha \approx 23{,}6; \quad \beta \approx 96{,}9,$$

$$f(x) = \frac{1}{\beta-\alpha} = \frac{1}{73{,}3} \approx 0{,}0136.$$



* Статистична перевірка гіпотез і критерії узгодження

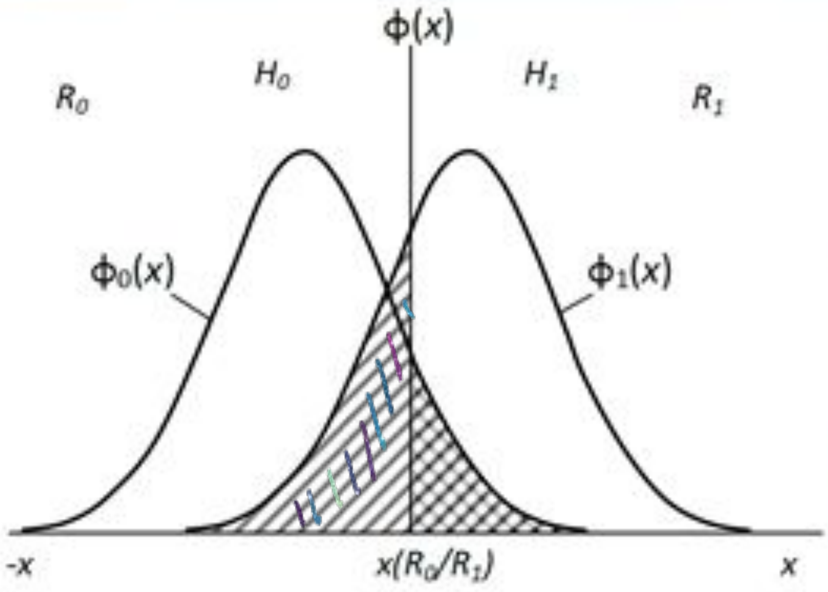
Однією з основних задач математичної статистики є визначення розподілу ймовірностей або параметрів цього розподілу за статистичними даними. При цьому задача іноді ставиться так.

Розглядають деяку гіпотезу про те, що розподіл ймовірностей має той чи інший вигляд, або параметри розподілу мають ті або інші значення. Задача полягає в тому, щоб на основі вивчення статистичних даних **підтвердити** правильність висунутої гіпотези **чи спростувати** її. Висунуту гіпотезу називають ***нульовою*** (***основною***) і позначають H_0.

Гіпотезу, що суперечить нульовій, називають ***конкуруючою*** (***альтернативною***) і позначають H_1.

Для перевірки різних гіпотез використовуються певні випадкові величини, які називають ***критеріями***.

* Статистичні гіпотези і їх перевірка



Густина ймовірності розподілу величини x

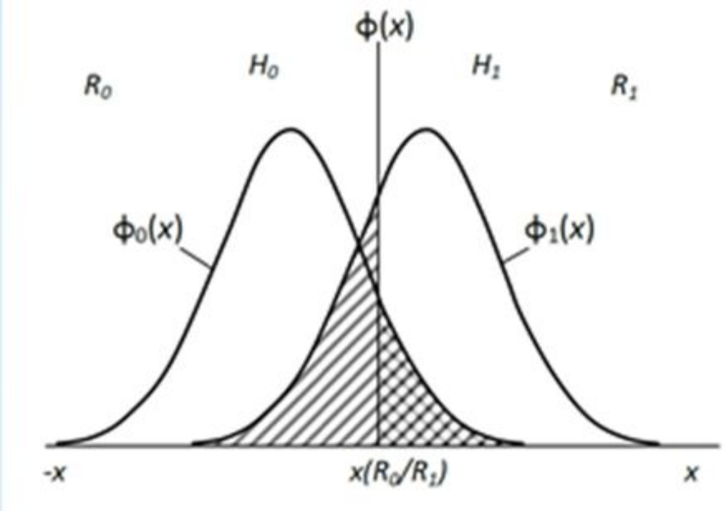
Критичною областю називають сукупність значень критерію, при яких нульову гіпотезу H0 відхиляють.

Критичні точки поділяють критичну область і область прийняття гіпотези. Критичних точок може бути одна чи дві, знаходять їх за спеціальними таблицями.

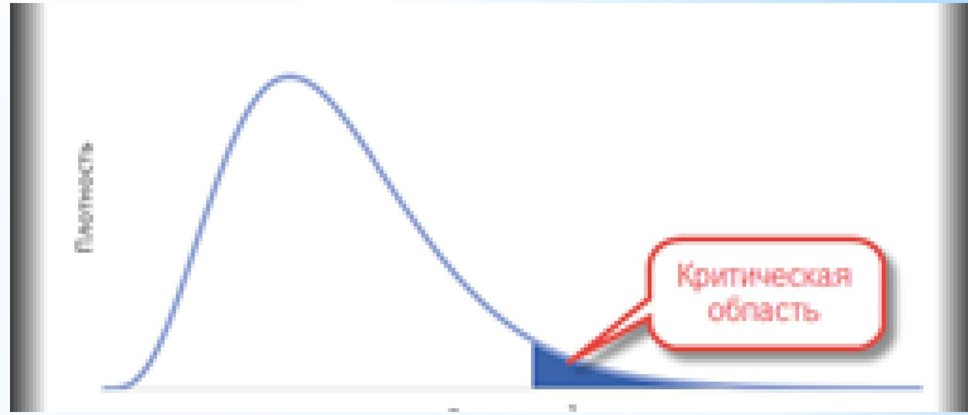
Рівнем значущості критерію перевірки статистичної гіпотези називається ймовірність того, що критерій не приведе до відхилення гіпотези в разі її істинності. Іншими словами, ***рівень значущості*** – це ймовірність того, що при істинності гіпотези спостережене значення не належатиме критичній області.

Рівень значущості α часто задають наперед, виходячи з особливостей задачі, і за рівнем значущості знаходять величину $U_{кр}$ так, щоб справджувалася рівність .

$$P(U \prec U_{кр}) = \alpha$$



Густина ймовірності розподілу величини *x*



* Якщо при заданому α виконується нерівність

$$U_{cn} \succ U_{kp}$$

то статистичні дані суперечать розглядуваній гіпотезі про вигляд розподілу ймовірностей, оскільки ймовірність

$$P(U_{cn} \geq U_{kp}) = 1 - \alpha$$

досить мала, коли α близьке до 1, а події, ймовірності яких не перевищують значення $1-\alpha$ вважаються практично неможливими. Імовірність того, що в цьому випадку буде відхилена правильна гіпотеза, дорівнює $1-\alpha$

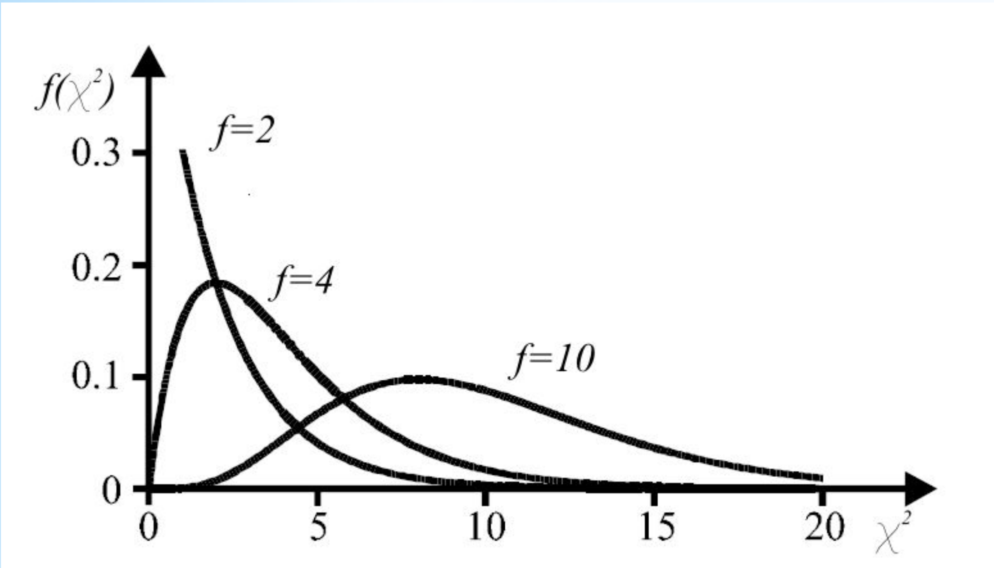
Якщо за статистичними даними дістанемо

$$U_{cn} \prec U_{kp}$$

то немає підстав відхиляти гіпотезу, оскільки статистичні дані не суперечать гіпотезі про те, що розподіл ймовірностей має саме той вигляд, про який йдеться в розглядуваній гіпотезі.

* Якщо спостережене значення U_{cn} належить критичній області, то гіпотеза відхиляється і приймається альтернативна гіпотеза, яка є запереченням основної гіпотези.

* Критерії узгодження (перевірки)
1. Критерій Пірсона;



$$k_r(u) = \begin{cases} \dfrac{1}{2^{\frac{r}{2}}\Gamma\left(\frac{r}{2}\right)} u^{\frac{r}{2}-1} e^{-\frac{u}{2}} & \text{при} \quad u > 0, \\ 0 & \text{при} \quad u < 0, \end{cases}$$

$$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t}\, dt \,.$$

χ^2 – критерій Пірсона

$$\chi^2 = n \sum_{i=1}^{k} \frac{(p_i^* - p_i)^2}{p_i} \quad ; \quad p_i^* = \frac{m_i}{n}$$

$$U = \chi^2 = \sum_{i=1}^{k} \frac{n^2\left(\frac{m_i}{n} - p_i\right)^2}{n p_i} = \sum_{i=1}^{k} \frac{n^2\left(\frac{m_i^2}{n^2} - 2\frac{m_i}{n}\cdot p_i + p_i^2\right)}{n\cdot p_i} =$$

$$= \sum_{i=1}^{k} \frac{m_i^2 - 2m_i n p_i + n^2 p_i^2}{n p_i} = \sum_{i=1}^{k} \frac{(m_i - np_i)^2}{np_i}$$

* Критерії погодження (перевірки)
1. Критерій Пірсона;

Таблица 4

Значения χ^2 в зависимости от r и p

r \ p	0,99	0,98	0,95	0,90	0,80	0,70	0,50	0,30	0,20	0,10	0,05	0,02	0,01	0,001
1	0,000	0,001	0,004	0,016	0,064	0,148	0,455	1,074	1,642	2,71	3,84	5,41	6,64	10,83
2	0,020	0,040	0,103	0,211	0,446	0,713	1,386	2,41	3,22	4,60	5,99	7,82	9,21	13,82
3	0,115	0,185	0,352	0,584	1,005	1,424	2,37	3,66	4,64	6,25	7,82	9,84	11,34	16,27
4	0,297	0,429	0,711	1,064	1,649	2,20	3,36	4,88	5,99	7,78	9,49	11,67	13,28	18,46
5	0,554	0,752	1,145	1,610	2,34	3,00	4,35	6,06	7,29	9,24	11,07	13,39	15,09	20,5
6	0,872	1,134	1,635	2,20	3,07	3,83	5,35	7,23	8,56	10,64	12,59	15,03	16,81	22,5
7	1,239	1,564	2,17	2,83	3,82	4,67	6,35	8,38	9,80	12,02	14,07	16,62	18,48	24,3
8	1,646	2,03	2,73	3,49	4,59	5,53	7,34	9,52	11,03	13,36	15,51	18,17	20,1	26,1
9	2,09	2,53	3,32	4,17	5,38	6,39	8,34	10,66	12,24	14,68	16,92	19,68	21,7	27,9
10	2,56	3,06	3,94	4,86	6,18	7,27	9,34	11,78	13,44	15,99	18,31	21,2	23,2	29,6
11	3,05	3,61	4,58	5,58	6,99	8,15	10,34	12,90	14,63	17,28	19,68	22,6	24,7	31,3
12	3,57	4,18	5,23	6,30	7,81	9,03	11,34	14,01	15,81	18,55	21,0	24,1	26,2	32,9
13	4,11	4,76	5,89	7,04	8,63	9,93	12,34	15,12	16,98	19,81	22,4	25,5	27,7	34,6
14	4,66	5,37	6,57	7,79	9,47	10,82	13,34	16,22	18.15	21,1	23,7	26,9	29,1	36,1
15	5,23	5,98	7,26	8,55	10,31	11,72	14,34	17,32	19,31	22,3	25,0	28,3	30,6	37,7
16	5,81	6,61	7,96	9,31	11,15	12,62	15,34	18,42	20,5	23,5	26,3	29,6	32,0	39,3
17	6,41	7,26	8,67	10,08	12,00	13,53	16,34	19,51	21,6	24,8	27,6	31,0	33,4	40,8
18	7,02	7,91	9,39	10,86	12,86	14,44	17,34	20,6	22,8	26,0	28,9	32,3	34,8	42,3
19	7,63	8,57	10,11	11,65	13,72	15,35	18,34	21,7	23,9	27,2	30,1	33,7	36,2	43,8
20	8,26	9,24	10,85	12,44	14,58	16,27	19,34	22,8	25,0	28,4	31,4	35,0	37,6	45,3
21	8,90	9,92	11,59	13,24	15,44	17,18	20,3	23,9	26,2	29,6	32,7	36,3	38,9	46,8
22	9,54	10,60	12,34	14,04	16,31	18,10	21,3	24,9	27,3	30,8	33,9	37,7	40,3	48,3
23	10,20	11,29	13,09	14,85	17,19	19,02	22,3	26,0	28,4	32,0	35,2	39,0	41,6	49,7
24	10,86	11,99	13,85	15,66	18,06	19,94	23,3	27,1	29,6	33,2	36,4	40,3	43,0	51,2
25	11,52	12,70	14,61	16,47	18,94	20,9	24,3	28,2	30,7	34,4	37,7	41,7	44,3	52,6
26	12,20	13,41	15,38	17,29	19,82	21,8	25,3	29,2	31,8	35,6	38,9	42,9	45,6	54,1
27	12,88	14,12	16,15	18,11	20,7	22,7	26,3	30,3	32,9	36,7	40,1	44,1	47,0	55,5
28	13,56	14,85	16,93	18,94	21,6	23,6	27,3	31,4	34,0	37,9	41,3	45,4	48,3	56,9
29	14,26	15,57	17,71	19,77	22,5	24,6	28,3	32,5	35,1	39,1	42,6	46,7	49,6	58,3
30	14,95	16,31	18,49	20,6	23,4	25,5	29,3	33,5	36,2	40,3	43,8	48,0	50,9	59,7

* Критерії погодження (перевірки)
1. Критерій Пірсона;

Критические точки распределения χ^2

Число степеней свободы k	Уровень значимости α					
	0,01	0,025	0,05	0,95	0,975	0,99
1	6,6	5,0	3,8	0,0039	0,00098	0,00016
2	9,2	7,4	6,0	0,103	0,051	0,020
3	11,3	9,4	7,8	0,352	0,216	0,115
4	13,3	11,1	9,5	0,711	0,484	0,297
5	15,1	12,8	11,1	1,15	0,831	0,554
6	16,8	14,4	12,6	1,64	1,24	0,872
7	18,5	16,0	14,1	2,17	1,69	1,24
8	20,1	17,5	15,5	2,73	2,18	1,65
9	21,7	19,0	16,9	3,33	2,70	2,09

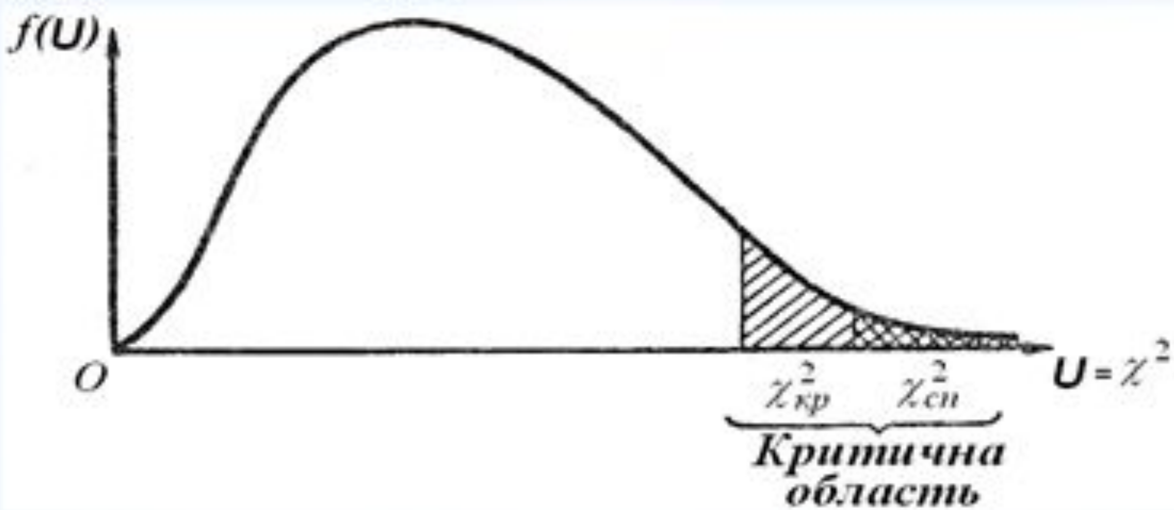
* **Для неперервного розподілу ймовірностей його застосовується так:**
* 1. За статистичними даними будують неперервний (інтервальний) розподіл частот, а на основі гіпотетичного розподілу ймовірностей знаходять гіпотетичні ймовірності попадання значень досліджуваної неперервної випадкової величини в часткові інтервали.
* Відхилення статистичного розподілу від гіпотетичного при цьому визначають за формулою

$$U_{cn} = \chi^2 = \sum_{i=1}^{k} \frac{(m_i - np_i)^2}{np_i}$$

Цей $\chi2$-розподіл не залежить від невідомого закону розподілу ймовірностей досліджуваної випадкової величини, а залежить лише від числа k степенів свободи.

- **Число ступенів свободи k** обчислюють за формулою $k = m - 1 - l$, де m – кількість часткових інтервалів, l – кількість параметрів гіпотетичного розподілу, які оцінюються за статистичними даними.
- Для $\chi2$-розподілу складені таблиці, за цими таблицями і підрахованим значенням k розраховують критичне значення $\chi^2_{кр}$

* коли $\chi^2_{сп} \geq \chi^2_{кр}$, то $\chi^2_{сп}$ належить критичній області (мал), а тому розглянуту гіпотезу про вид розподілу ймовірностей **слід відхилити**, оскільки вона при даному рівні значимості не узгоджується зі статистичними даними.



• Якщо $\chi^2_{сп} \prec \chi^2_{кр}$, то вважають, що розглядувана гіпотеза не суперечить статистичним даним.

Задача 2. Проведено 500 вимірювань бокової похибки наводки при стрільбі з літака по наземній цілі. Результати вимірювань (в тисячних долях радіана) зведено в статистичний ряд. Побудувати гістограму, наближену статистичну функцію розподілу за даними статистичного ряду.

I_i	−4; −3	−3; −2	−2; −1	−1; 0	0, 1	1; 2	2; 3	3; 4
m_i	6	25	72	133	120	88	46	10
p_i^*	0,012	0,050	0,144	0,266	0,240	0,176	0,092	0,020

* Вирівнювання статистичного розподілу задачі 2 за допомогою нормального закону.

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}$$

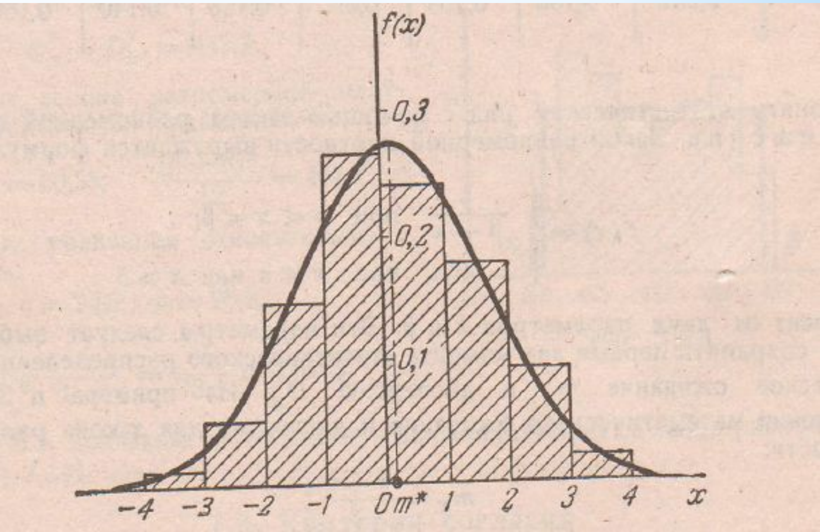
$$m_x^* = -3{,}5\cdot 0{,}012 - 2{,}5\cdot 0{,}050 - 1{,}5\cdot 0{,}144 - 0{,}5\cdot 0{,}266 + 0{,}5\cdot 0{,}240 + 1{,}5\cdot 0{,}176 + 2{,}5\cdot 0{,}092 + 3{,}5\cdot 0{,}020 = 0{,}168.$$

$$\alpha_2^* = \sum_{i=1}^{\infty} \tilde{x}_i^2 p_i^* = 2{,}126.$$

$$D_x^* = \alpha_2^* - (m_x^*)^2 = 2{,}126 - 0{,}028 = 2{,}098.$$

$$m = 0{,}168; \quad \sigma = 1{,}448.$$

$$f(x) = \frac{1}{1{,}448\sqrt{2\pi}} e^{-\frac{(x-0{,}168)^2}{2\cdot 1{,}448^2}}$$



x	—4	—3	—2	—1	0	1	2	3	4
$f(x)$	0,004	0,025	0,090	0,199	0,274	0,234	0,124	0,041	0,008

* **Для задачі №2 перевірити узгодженість теоретичного і статистичного розподілень.**

$m = 0{,}168, \qquad \sigma = 1{,}448,$

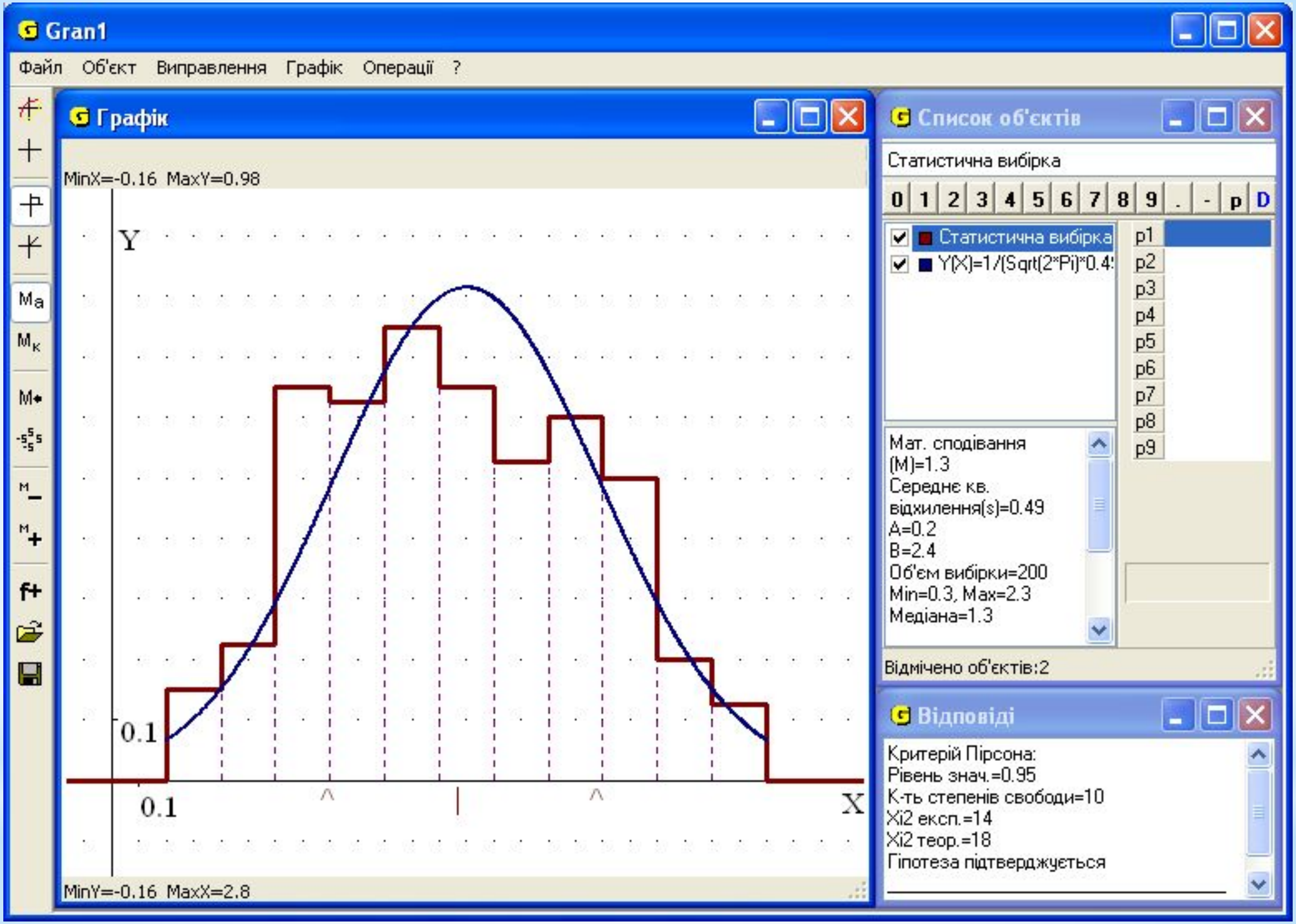
$$p_i = \Phi^*\left(\frac{x_{i+1} - m}{\sigma}\right) - \Phi^*\left(\frac{x_i - m}{\sigma}\right),$$

I_i	—4; —3	—3; —2	—2; —1	—1; 0	0; 1	1; 2	2; 3	3; 4
m_i	6	25	72	133	120	88	46	10
np_i	6,2	26,2	71,2	122,2	131,8	90,5	38,2	10,5

$$\chi^2 = \sum_{i=1}^{8} \frac{(m_i - np_i)^2}{np_i} = 3{,}94.$$

$r = 8 - 3 = 5.$

при $\chi^2 = 3{,}00 \quad p = 0{,}70;$
при $\chi^2 = 4{,}35 \quad p = 0{,}50.$


Gran1
Файл Об'єкт Виправлення Графік Операції ?
Графік
MinX=-0.16 MaxY=0.98
Y
0.1
0.1
X
MinY=-0.16 MaxX=2.8
Список об'єктів
Статистична вибірка
0 1 2 3 4 5 6 7 8 9 . - p D
Статистична вибірка
Y(X)=1/(Sqrt(2*Pi)*0.4!
p1
p2
p3
p4
p5
p6
p7
p8
p9
Мат. сподівання
(M)=1.3
Середнє кв.
відхилення(s)=0.49
A=0.2
B=2.4
Об'єм вибірки=200
Min=0.3, Max=2.3
Медіана=1.3
Відмічено об'єктів:2
Відповіді
Критерій Пірсона:
Рівень знач.=0.95
К-ть степенів свободи=10
Xi2 експ.=14
Xi2 теор.=18
Гіпотеза підтверджується

* **Задача. З метою дослідження закону розподілу похибки вимірювання дальності за допомогою радіодальноміра зроблено 400 вимірювань. Результати досліду представлені у вигляді статистичного ряду. Вирівняти ряд за допомогою закону рівномірної густини**

I_i (м)	20; 30	30; 40	40; 50	50; 60	60; 70	70; 80	80; 90	90; 100
m_i	21	72	66	38	51	56	64	32
p_i^*	0,052	0,180	0,165	0,095	0,128	0,140	0,160	0,080

$$f(x) = \begin{cases} \dfrac{1}{\beta-\alpha} & \text{при} \quad \alpha < x < \beta; \\ 0 & \text{при} \quad x < \alpha \text{ или } x > \beta \end{cases}$$

$$m_x = \frac{\alpha+\beta}{2};$$

$$D_x = \frac{(\beta-\alpha)^2}{12}.$$

$\tilde{x}_i'$	—35	—25	—15	—5	5	15	25	35
p_i^*	0,052	0,180	0,165	0,095	0,128	0,140	0,160	0,080

$$m^*_{x'} = \sum_{i=1}^{k} \tilde{x}'_i p^*_i = 0{,}26.$$

$$\alpha^*_2 = \sum_{i=1}^{k} \left(\tilde{x}'_i\right)^2 p^*_i = 447{,}8,$$

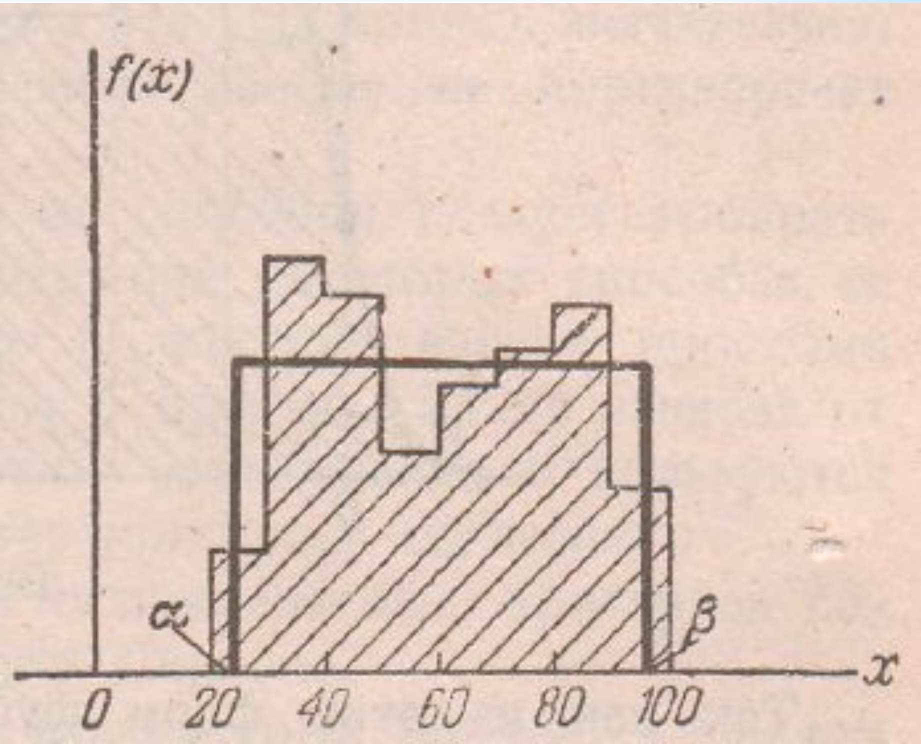
$$D^*_{x'} = \alpha^*_2 - \left(m^*_{x'}\right)^2 = 447{,}7.$$

$$m^*_x = m^*_{x'} + 60 = 60{,}26$$

$$\frac{\alpha + \beta}{2} = 60{,}26; \qquad \frac{(\beta - \alpha)^2}{12} = 447{,}7.$$

$$\alpha \approx 23{,}6; \quad \beta \approx 96{,}9,$$

$$f(x) = \frac{1}{\beta - \alpha} = \frac{1}{73{,}3} \approx 0{,}0136.$$



J_i	20;30	30;40	40;50	50;60	60;70	70;80	80;90	90;100
m_i	21	72	66	38	51	56	64	32
$n \cdot p_i$	34,9	54,6	54,6	54,6	54,6	54,6	54,6	38

$$f(x) = \frac{1}{73{,}3}$$

$$P(x) = \frac{\beta - \alpha}{73{,}3}$$

$$\left.\begin{array}{l} \alpha = 23{,}6 \\ \beta = 96{,}9 \end{array}\right\} \Rightarrow \begin{array}{l} np_1 = \frac{30 - 23{,}6}{73{,}3} \cdot 400 = 34{,}9 \\ np_8 = \frac{96{,}9 - 90}{73{,}3} \cdot 400 = 38 \end{array}$$

$$\chi^2 = \sum_{i=1}^{8} \frac{(m_i - np_i)^2}{np_i} \approx 20{,}9$$

$$f = 8 - 1 - 2 = 5$$

$$\chi^2_{кр}(5,\ 0{,}95) = 11{,}1$$

Гіпотеза не приймається

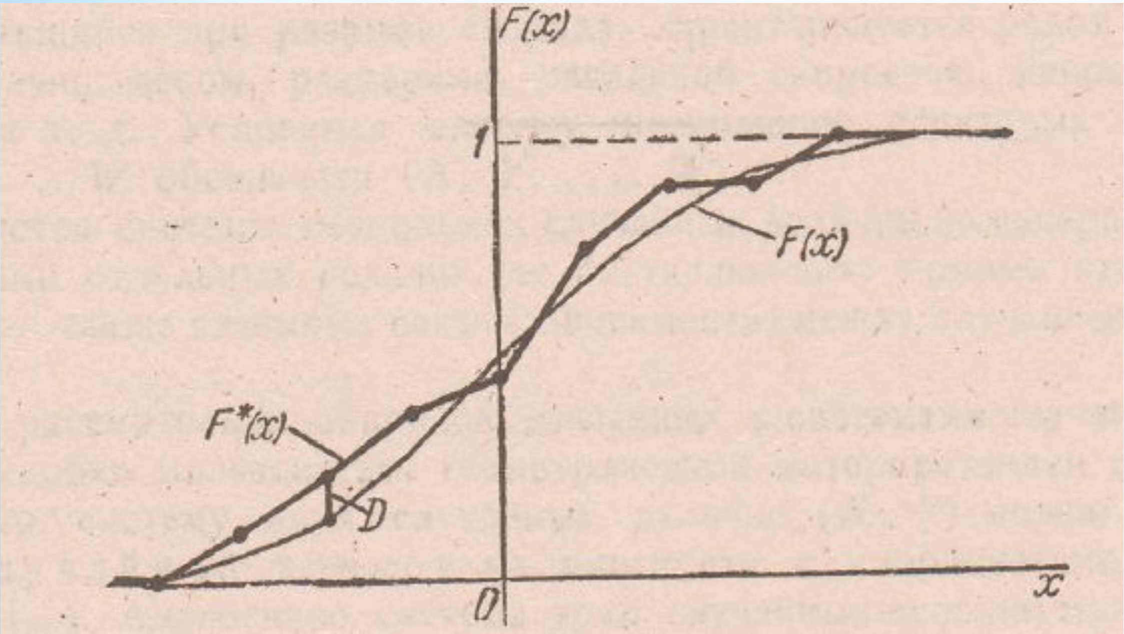
* Критерій Колмогорова

$$D = \max |F^*(x) - F(x)|.$$

$$D\sqrt{n} \geqslant \lambda$$

$$P(\lambda) = 1 - \sum_{k=-\infty}^{\infty} (-1)^k e^{-2k^2\lambda^2}.$$

λ	$P(\lambda)$	λ	$P(\lambda)$	λ	$P(\lambda)$
0,0	1,000	0,7	0,711	1,4	0,040
0,1	1,000	0,8	0,544	1,5	0,022
0,2	1,000	0,9	0,393	1,6	0,012
0,3	1,000	1,0	0,270	1,7	0,006
0,4	0,997	1,1	0,178	1,8	0,003
0,5	0,964	1,2	0,112	1,9	0,002
0,6	0,864	1,3	0,068	2,0	0,001



Обмеження для коректного застосування критерію:

- Вибірка досить велика (більше 50 спостережень).
- Класи інтервалів повинні бути впорядковані по зростанню або зменшенню деякої ознаки. Вони обов'язково повинні відображати її направлену зміну.

Практичне застосування критерію Колмагорова

1. Формуємо простий варіаційний ряд:

$$x_1 \leq x_2 \leq \ldots \leq x_n$$

2. Обчислення різниці $\frac{i}{n} - F(x_i)$ для $i = 1 \ldots n$ і найбільше значення (за абсолютною величиною), позначаємо d^+:

$$d^+ = \max_{1 \leq i \leq n} \left| \frac{i}{n} - F(x_i) \right|$$

3. Обчислюємо всі різниці

$$F(x_i) - \frac{i-1}{n}$$

і найбільше за абсолютною величеною позначаємо d^-:

$$d^- = \max_{1 \leq i \leq n} \left| F(x_i) - \frac{i-1}{n} \right|$$

4. Вибираємо більше з чисел d^+, d^-

$$d = \max(d^+, d^-)$$

5. За даним рівнем значущості α та n визначаємо з таблиці критичне значення $d_{кр}$ статистики Колмагорова D.

Якщо обчислене значення $d < d_{кр}$, то висунута гіпотеза приймається.

Критичні точки $d_n(1-\alpha)$ статистики D_n Колмогорова

n	α			n	α		
	0,10	0,05	0,01		0,10	0,05	0,01
1	0,950	0,975	0,995	26	,233	,259	,311
2	,776	,842	,929	27	,229	,254	,305
3	,636	,708	,829	28	,225	,250	,300
4	,565	,624	,734	29	,221	,246	,294
5	,509	,563	,669	30	,218	,242	,290
6	,468	,519	,617	31	,214	,238	,285
7	,436	,483	,576	32	,211	,234	,281
8	,410	,454	,542	33	,208	,231	,277
9	,387	,403	,513	34	,205	,227	,273
10	,369	,409	,489	35	,202	,224	,269
11	,352	,391	,468	36	,199	,221	,265
12	,338	,375	,449	37	,196	,218	,262
13	,325	,361	,432	38	,194	,215	,258
14	,314	,349	,418	39	,191	,213	,255
15	,304	,338	,404	40	,189	,210	,252
16	,295	,327	,392	41	,187	,208	,249
17	,286	,318	,381	42	,185	,205	,246
18	,279	,309	,371	43	,183	,203	,243
19	,271	,301	,361	44	,181	,201	,241
20	,265	,294	,352	45	,179	,198	,238
21	,259	,287	,344	46	,177	,196	,235
22	,253	,281	,337	47	,175	,194	,233
23	,247	,275	,330	48	,173	,192	,231
24	,242	,269	,323	49	,171	,190	,228
25	,238	,264	,317	50	,170	,188	,226

Критерій узгодження Колмогорова.

*

i	x_i	$\frac{i}{n}$	$F(x_i)$	$\frac{i-1}{n}$	$\left\lvert\frac{i}{n} - F(x_i)\right\rvert$	$\left\lvert\frac{i-1}{n} - F(x_i)\right\rvert$
1	0,18	0,2	0,1647	0	0,0353	0,1647
2	0,56	0,4	0,4288	0,2	0,0288	0,2288
3	0,87	0,6	0,5810	0,4	0,0190	0,1810
4	1,37	0,8	0,7459	0,6	0,0541	0,1459
5	2,46	1	0,9146	0,8	0,0854	0,1146

За рівня значущості $\alpha = 0{,}05$ за допомогою критерія Колмогорова перевірити гіпотезу H_0, що вибірку взято з генеральної сукупності, в якій $F(x)$:

$$F(x) = \begin{cases} 0, & x < 0 \\ 1 - e^{-x}, & x \geq 0 \end{cases}$$

$$d_{кр} = d_5(1-\alpha) = d_5(0{,}95) = 0{,}563$$

Гіпотеза приймається